



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :  C12Q 1/68, C12N 9/12	A1	(11) International Publication Number: WO 90/08839  (43) International Publication Date: 9 August 1990 (09.08.90)
(21) International Application Number: PCT/US90/00440  (22) International Filing Date: 31 January 1990 (31.01.90)  (30) Priority data: 306,423 6 February 1989 (06.02.89) US  (71) Applicant: EASTMAN KODAK COMPANY [US/US]; 343 State Street, Rochester, NY 14650 (US).  (72) Inventors: VIZARD, Douglas, Lincoln ; 663 Meriden Road, Cheshire, CT 06410 (US). HADMAN, Martin ; 74 Liberty Street, Stamford, CT 06902 (US).  (74) Agent: EVERETT, John, R.; 343 State Street, Rochester, NY 14650 (US).	(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (Eu- ropean patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European pa- tent), SE (European patent).  Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	

(54) Title: THERMOSTABLE POLYMERASE DNA SEQUENCING REACTION CONCENTRATE

(57) Abstract

The present invention provides a nucleotide sequencing reaction concentrate comprising: a) sufficient thermostable polymerase to provide from 100 to 500 I.U./mL; b) from 0.3 to 30 mM of a reducing agent selected from dithiothreitol and  $\beta$ -mercaptoethanol; c) sufficient phosphate buffer to maintain a pH of about 7.5; d) at least 40% glycerol; e) from 15 to 70  $\mu$ M dGTP or daGTP; f) from 15 to 150  $\mu$ M dATP; g) from 15 to 150  $\mu$ M dTTP; h) from 10 to 18  $\mu$ M dCTP; and i) one member selected from the group consisting of: i) sufficient ddGTP to provide a ddGTP to dGTP or daGTP ratio of about 5; ii) sufficient ddATP to form a ddATP to dATP ratio of about 33; iii) ddTTP in a sufficient amount to form a ddTTP to dTTP ratio of about 28; and iv) ddCTP in a sufficient amount to form a ddCTP to dCTP ratio of about 22.

-1-

THERMOSTABLE POLYMERASE DNA SEQUENCING  
REACTION CONCENTRATE

Field of the Invention

This invention relates to DNA sequencing, a reaction concentrate for sequencing, a primer for the sequencing and a kit comprising the reaction concentrate.

BACKGROUND OF THE INVENTION

There are several method for sequencing DNA. One of the most frequently used is that of Sanger et al reported in Proc. Natl. Acad. Sci. USA volume 74, No. 12, pp. 5463-5467. Sanger's method is based on the inhibitory activity of nucleotide analogues such as dideoxy nucleotide triphosphates (ddNTPs where N is thymine, guanine, cytosine or adenine) on the ability of DNA polymerase to catalyze the growth of a oligonucleotide chain.

For example, when a primer and template are incubated with DNA polymerase in the presence of a mixture of ddTTP and dTTP, as well as the other three deoxyribonucleosides triphosphates (dNTP wherein N is G, A and C) one of which is labeled with a radioactive nucleotide, a mixture of fragments will be synthesized all having the same 5' end with ddT residues at the 3' end is obtained. When this mixture is fractionated by electrophoresis on denaturing acrylamide gels, the pattern of bands show the distribution of dTs in the newly synthesized DNA. By using analogous terminators for other nucleotides in separate incubations and analyzing the samples in parallel on a gel, a pattern of bands is obtained from which the sequence can be deduced.

In general, a DNA fragment to be sequenced is spliced into a particular site in the DNA of a sequencing vector, such as M13 phage, using standard recombinant DNA splicing methods. The resulting

-2-

recombinant phage DNA is used to infect bacterial cells, and virus particles are produced. These particles contain only one of two strands of DNA.

The single strands of DNA are purified from 5 the viruses. They are mixed with pieces of DNA referred to as primers. Primers are complementary to a region of the M13 DNA near the position where the DNA to be sequenced has been inserted. The primer forms base pairs with the M13 DNA creating a short 10 double-stranded region of DNA that serves as a primer for synthesis. A large number of these partially double-stranded DNA molecules are mixed with DNA polymerase and a radioactive dNTP. DNA polymerase synthesizes radioactive DNA from the M13 template 15 (beginning at the 3' end of the short double-stranded region). The synthesis ultimately involves the DNA region that is to be sequenced and which acts as a template.

The synthesis can be terminated by including 20 a ddNTP in the biochemical reaction. If the ddNTP is ddTTP, the new radioactive chain will stop after it would normally have added a T. Normal dTTP is included in the reaction mix to allow some DNA synthesis before a ddTTP stops the synthesis. Since 25 there are many T's in a nucleotide sequence the stops will not always be in the same place. Thus by carefully adjusting the amount of analog in the reaction it is possible to create a collection of radioactive DNA fragments that begin at the 3' end of 30 the primer and stop at the various positions where T's occur. This procedure is repeated in three other test tubes comprising all of the necessary dNTPs and an analog ddGTP, ddCTP or ddATP. This results in four separate collections of DNA fragments. Every 35 fragment in each tube starts at the same place and ends at different distances in a G, A, T or C dependent on the specific ddNTP included.

-3-

The lengths of the molecules are measured by gel electrophoresis under conditions that are so precise that they can separate DNA sequences differing by only one nucleotide. The four separate 5 collections are electrophoresed next to each other and after electrophoresis a photographic film is exposed to the radioactive emissions of the DNA in the prepared electrophoretic gel. By analysis of the developed film one is able to read the DNA sequence 10 directly from the photographic film.

This method normally involves many different steps that must be carefully carried out in order to achieve precise and accurate sequence analyses. The elimination of any of the steps would improve the 15 method. For example, the method as generally performed requires separate mixing of all the chemical components involved at the time of use. That is the buffer, enzyme, nucleotides, nucleotide analogues and oligonucleotide primer are mixed 20 together separately for each nucleotide G, T, A and C.

#### SUMMARY OF THE INVENTION

The present invention provides a nucleotide sequencing reaction concentrate comprising:

- a) sufficient thermostable polymerase to 25 provide from 100 to 500 I.U./mL;
- b) from 0.3 to 30 mM of a reducing agent selected from dithiothreitol and  $\beta$ -mercaptoethanol;
- c) sufficient phosphate buffer to maintain a pH of about 7.5;
- d) at least 40% glycerol;
- e) from 15 to 150  $\mu$ M dGTP or daGTP (where "a" 30 G" represents 7-deazaguanine);
- f) from 15 to 150  $\mu$ M dATP;
- g) from 15 to 150  $\mu$ M dTTP;
- h) from 10 to 18  $\mu$ M dCTP; and
- i) one member selected from the group consisting of:

-4-

- i) sufficient ddGTP to provide a ddGTP to dGTP or daGTP ratio of about 5;
- ii) sufficient ddATP to form a ddATP to dATP ratio of about 33;
- 5           iii) ddTTP in a sufficient amount to form a ddTTP to dTTP ratio of about 28; and
- iv) ddCTP in a sufficient amount to form a ddCTP to dCTP ratio of about 22.

Use of the reaction concentrate of this

- 10 invention eliminates the need to mix all the components needed for sequencing a nucleotide chain at the time of use. The reaction concentrate comprises all of the essential components in a premixed package that can be stored and reused during
- 15 several cycles. One advantage of the reaction concentrates is the simplicity of the method involving execution of the sequencing reaction steps. Such steps involve the fewest number of steps of all current sequencing procedures and approaches a
- 20 desirable "single-step" protocol for sequencing.

In a preferred embodiment the reaction concentrate comprises, in addition to the components listed above:

- a) from 175 to 700 $\mu$ g/mL of concentrate of a
- 25 protein selected from the group consisting of gelatin and bovine serum albumin;
- b) at least 25 mM of potassium glutamate;
- c) from 0.1% to 0.35% V/V of a non-ionic surfactant; and
- 30       d) from 9 to 35 mM of a magnesium salt such as magnesium acetate or magnesium chloride.

In this embodiment the reaction concentrate can be stored at temperatures as low as -80°C to -20°C without any deleterious effect on the ability 35 of the components to react in the chemical reactions involved in nucleotide sequencing.

-5-

In practice, the dCTP concentration in the reaction can be reduced to 3 $\mu$ M and still support sufficient enzyme reaction kinetics. This is the only one of the four nucleotides that can be reduced 5 to that level. Consequently, dCTP is the recommended target nucleotide used for radioactive labeling of the synthesized strand. Such a low concentration of the target nucleotide is essential for efficient labeling, especially if the  $^{35}$ S thioester of dCTP 10 is to be used. Other radioactive nucleotides can be used with substantially reduced efficiency.

DETAILS OF THE INVENTION

The essential components of the reaction concentrates are 100 to 500, preferably 250 I.U./mL 15 of a thermostable polymerase, preferably Thermus aquaticus polymerase (Taq pol); dithiothreitol (DTT); phosphate buffer; magnesium salt; glycerol; the deoxynucleotide triphosphates dNTPs wherein N is guanine (G), 7-deazaguanine (aG), adenine (A), 20 thymine (T) and cytosine (C); and a ddNTP (dideoxy NTP wherein N is as previously defined) to terminate the G, A, T and C reactions.

In one embodiment this invention provides four identical reaction concentrates differing only 25 in the concentration of each dNTP, the specific ddNTP and the ratio of ddNTP to dNTP. The specific ddNTP would depend on which nucleotide is designated as the terminating nucleotide in any polymeric sequence. According to the Sanger method that would be G, A, T 30 or C.

In addition to the foregoing essential elements, reaction concentrate can optimally contain a) a protein; b) a non-ionic detergent; and c) potassium glutamate.

35 The reaction concentrate can be made even more convenient to the ultimate user by including

-6-

therein a primer as well as the radioactive labeled dNTP. However, the short half life of radioactive dNTPs makes the inclusion thereof impractical.

Reaction concentrates can be constructed  
5 that contain nearly all of the essential components to DNA sequencing. This presents the opportunity for a user to simply add a solution of DNA to be sequenced to a mixture of reaction concentrate, incubate the mixture at elevated temperature.  
10 terminate and analyze the reaction products as in contemporary practice. Such method eliminates many steps and is far more amenable to automation than any contemporary sequencing methodology.

The thermostable enzyme is present in the  
15 concentrate in sufficient amount to provide an activity of about 0.5 I.U per base-specific reaction. I.U. is the amount of enzyme required to catalyze a polymerization of 10 micromoles of substrate nucleotide per 30 minutes under a broad pH  
20 range and a temperature of 70°C for this enzyme.

As used herein, the term "thermostable enzyme" refers to an enzyme which is stable to heat and is heat resistant and catalyzes polymerization of nucleotides to form a nucleotide sequence that is  
25 complementary to the unknown nucleotide sequence. Generally, the synthesis will be initiated at the 3' end of each primer and will proceed in the 5' to 3' direction along the template strand thereby proceeding into the unknown nucleotide sequence of the  
30 sequencing vector, until synthesis is terminated by a ddNTP, producing oligonucleotides of different lengths. There may be a thermostable enzyme, however, which initiates synthesis at the 5' end and proceeds in the other direction, using the same  
35 process as described above.

-7-

The preferred thermostable enzyme herein is a DNA polymerase isolated from Thermus aquaticus (Taq pol). Various strains thereof are available from the American Type Culture Collection, Rockville, Maryland, and is described by T.D. Brock, J. Bact. (1969) 98:289-297, and by T. Oshima, Arch. Microbiol. (1978) 117:189-196. One of these preferred strains is strain YT-1.

Our experiments have shown that Taq Pol

10 1. does not react at low temperatures and can be stored in elevated salt and appropriate reaction buffer conditions at -20°C;

15 2. is relatively free of contaminating nucleases and phosphatases; and

15 3. maintains appreciable activity in relatively high concentrations of glycerol. These observations made the present reaction concentrates possible.

The thermostable enzyme may also be produced by recombinant DNA techniques, as the gene encoding this enzyme has been cloned from Thermus aquaticus genomic DNA. The complete coding sequence for the Thermus aquaticus (Taq) polymerase can be derived from bacteriophage CH35:Taq#4-2 on an approximately 3.5 kilobase (kb) BglIII-Asp<sup>718</sup> partial restriction

25 fragment contained within an ~18 kb genomic DNA insert fragment. This bacteriophage was deposited with the American Type Culture Collection (ATCC) on May 29, 1987 and has accession no. 40,336.

Alternatively, the gene can be constructed by

30 ligating an ~750 base pair (bp) BglIII-HinIII restriction fragment isolated from plasmid pFC83 (ATCC 67,422, deposited May 29, 1987) to an ~2.8 kb HinIII-Asp<sup>718</sup> restriction fragment isolated from plasmid pFC85 (ATCC 67,421, deposited May 29, 1987).

35 The pFC83 restriction fragment comprises the amino-terminus of the Taq polymerase gene while the

-8-

restriction fragment from pFC85 comprises the carboxyl-terminus. Thus, ligation of these two fragments into a correspondingly digested vector with appropriate control sequences will result in the 5 translation of a full-length Taq polymerase.

These parameters were determined experimentally. Less than the 0.5 I.U. per base-specific reaction results in artificial termination of nucleotide chains. Greater than 2 I.U. per 10 base-specific reaction is unnecessary.

The protein can be bovine serum albumin (BSA) or gelatin. Taq pol enzymatic activity is improved by a protein environment. The protein also enhances the shelf life of the reaction concentrate.

15 The presence of the protein also seems to avoid the inclusion of artifacts in a polymerized nucleotide sequence in which Taq pol catalyzed polymerizations amounts of 175-700 $\mu$ g/mL of protein in the reaction concentrate, preferably 350 $\mu$ g/mL, are desirable.  
20 Concentrations below 175 $\mu$ g/mL result in instability. Concentrations greater than 700 $\mu$ g/mL are unnecessary.

A reducing agent is essential.

Dithiothreitol (DTT) is preferred. It maintains the 25 sulfhydryl groups of Taq pol in a reduced state. This reduced state is necessary for Taq pol activity. Other reducing agents include  $\beta$ -mercaptoethanol. Amounts from 0.3 to 10 mM of DTT are useful, preferably 3.5 mM. Concentrations 30 outside of this range result in diminished rates of polymerization. Concentrations above this range are unnecessary.

Commercial preparation of Taq pol usually contains small amounts (about 0.1 mM) of EDTA as an 35 aid for preservative purposes. However, more than 1mM EDTA would be detrimental to the reaction

- 6 -

concentrate because EDTA is a metal scavenger and would therefore have an affect on magnesium ion activity which is essential for Taq pol activity.

The reaction concentrate preferably contains  
5 up to 17.5 mM of a magnesium salt such as magnesium chloride or magnesium acetate ( $\pm 50\%$ ). Magnesium acetate is preferred. It is essential that it is present at about a 5 mM concentration in the final reaction mix. Amounts in the ranges of  $\pm 50\%$  of  
10 5 mM magnesium acetate will be effective. Above or below these ranges the magnesium will have a deleterious affect on the enzyme activity.

Potassium glutamate (K glutamate) also enhances Taq pol enzymatic activity. It is essential  
15 for long term storage of the reaction concentrate. Long term storage requires a minimum amount of 25 mM of the potassium glutamate. Higher concentrations lead to an unacceptable salt concentration that would have a deleterious affect on the electrophoretic  
20 analysis carried out in the sequencing method provided by this invention.

Potassium hydrogen phosphate and potassium dihydrogen phosphate ( $K_2HPO_4$  and  $KH_2PO_4$ ) should be present in an amount sufficient to  
25 establish a pH of about 7.5. Phosphate buffer is a thermally stable buffer. It is able to maintain essentially the same pH over a temperature range of -20° to 75°C. It is necessary to maintain the pH constant in storage as well as in use so that Taq pol  
30 activity is constant from storage to use.

Glycerol is an antifreeze. Repeated freezing of any enzyme must be prevented to maintain enzyme activity as the reaction concentrate goes through the expected cycles of thawing for use and  
35 then restorage for future uses. We have found that at storage temperatures of -70°C to -20°C at least

-10-

40% glycerol must be added to the reaction concentrate to prevent repeated freezing. Up to 80% of glycerol could be used without adversely affecting Taq pol activity.

5       The reaction concentrate can comprise from 0.1 to 0.4% of a non-ionic detergent. Examples include ethoxylated fatty alcohol ethers and lauryl ethers, ethoxylated alkyl phenols, octylphenoxy polyethoxy ethanol compounds, modified oxyethylated  
10 and/or oxypropylated straight-chain alcohols, polyethylene glycol monooleate compounds, polysorbate compounds, and phenolic fatty alcohol ethers. More particularly preferred are TRITON X-100, Tween 20, from ICI Americas Inc., Wilmington, DE, which is a  
15 polyoxyethylated (20) sorbitan monolaurate, and Iconol<sup>TM</sup> NP-40, from BASF Wyandotte Corp., Parsippany, NJ, which is an ethoxylated alkyl phenol(nonyl).

Each of the four dNTPs are also essential  
20 components of each concentrate. Each of the four concentrates must contain one, and only one, of the ddNTPs according to the termination reaction designed into each reaction concentrate package.

In one embodiment the materials of the  
25 present invention can be provided to the ultimate user as a kit in several embodiments. Consistent with the spirit of this invention of user convenience, it is desirable to provide a kit in which the kit components can be used directly without dilution  
30 steps. Thus the most convenient kit would be designed to allow the user, based on the practice in many biochemistry labs, to take approximately 2 $\mu$ L of the reaction concentrate for direct use.

In one embodiment the kit can comprise  
35 distinct and separately packaged chemicals which may be labeled 1) primer and 2) separate reaction concentrates G, A, T and C.

-11-

An optional user variable package can include a sequencing vector or a control DNA and a radioactive labeled dNTP. The control has a known DNA sequence which allows the user to confirm that the  
5 kit is operative.

The sequencing vector comprises the nucleotide chain to be sequenced. The most widely used sequencing vectors have been derived from bacteriophage M13. This vector is widely available  
10 commercially in various formats usually differing only by their multiple cloning regions. Widely used versions of this vector include M13mp18 and mp19. Other commercially available sequencing vector templates include pUC, pIBI, Bluescript and others.  
15 If the primer described hereinafter is used, any sequencing vector having a lacZ nucleotide sequence can be used.

Alternatively, the user may choose to supply the primer and not use the supplied primer. The user  
20 may also choose to independently label the primer with a radioactive or fluorescent label. In this situation, additional use of radioactive labels is unnecessary. As long as the primer concentration and binding ability is substantially unmodified, the  
25 reaction concentrations of this invention will be useful.

Commercially supplied radioactive nucleotides are sold at a concentration range of about 3 to 16  $\mu\text{M}$  in water. This amount (about 1  $\mu\text{L}$ ) is added  
30 as a user variable where that amount is then further diluted by adding a certain amount of the user variable part of the composition to each of 4 different reaction concentrates. That is, the 1  $\mu\text{L}$  of radioactivity as a user variable is co-mixed with  
35 the user's DNA or control DNA, water and 5  $\mu\text{L}$  of primer to total 20 to 23  $\mu\text{L}$ , and is divided 4 times

-12-

to end up with 2.5  $\mu$ Ci per 5  $\mu$ L to which 2  $\mu$ L of reaction concentrate are added, giving four final reaction mixtures of 7  $\mu$ L each.

The required amount of radionucleotide for 5 each of the 4 reactions is based on completing the sequence analysis in 24 hours with use of a preferred isotope. The preferred isotope is alpha-labeled  $^{35}$ S dCTP although others such as  $^{32}$ P could be used. The preferred range would therefore be 1-2  $\mu$ L 10 of alpha-labeled  $^{35}$ S dCTP in the kit. This provides about 10  $\mu$ Ci/ $\mu$ L at a specific activity of 400-600 Ci per mM of  $^{35}$ S dCTP.

The user can supply the radioactive dNTP or it can be included in each of the reaction 15 concentrates. When the user provides both the sequencing vector and the radioactive dNTP, this constitutes the user variables.

One version of a user variable is presented below. These variables are supplied by the user. 20 Optionally they may be supplied as part of a kit.

	Concentration in Final Reaction
	<u>Mixture</u>
<u>User Variable Concentrate</u>	
1 $\mu$ g M13mp18 or a user selected	
25 sequencing vector amounting to	
0.4 pmoles of the	
DNA to be sequenced.....	0.1 pmoles/7 $\mu$ L
1 $\mu$ L alpha-labeled dCTP.....	$2.5\mu$ Ci/7 $\mu$ L
Plus water to a final	
30 volume not to exceed 18 $\mu$ L	

The primer concentrate package comprises an appropriate concentration of an oligonucleotide primer in a volume of 5  $\mu$ L.

35 Essential to Taq pol sequencing is the extension of an appropriate primer. Elevated

-13-

temperature of synthesis does not permit effective use of "universal primer A", which is a 17 base oligomer described herein below. Using longer oligomers does permit "spontaneous" priming of template DNA at elevated temperatures, and using such an appropriate primer obviates the need for the priming reaction essential to other DNA sequencing methodologies. The phrase "spontaneous" means that the priming reaction occurs without intervention by the user.

In a kit according to this invention the primer includes the commonly used universal primer region of M13mp18, but also includes additional bases in the surrounding sequence, taking advantage of more stable G and C residues. This oligomer is compatible with all lacZ based vectors (M13, pUC, pIBI, Bluescript and others). The oligomer sequence is:

5' CCCAGTCCACGACGTTGTAAAACGACGGCCAGTG 3'

The underlined region is the traditionally used universal primer A region. If primers shorter than 33 mer are used, excessive molar amounts are required for spontaneous priming. Such excessive amounts will result in non-specific priming leading to ambiguous sequencing results.

Longer primers can also be used. However, they seriously compromise sequencing results. Extending in the 5' direction leads to G rich sequence inclusions, resulting in ambiguous electrophoretic mobilities caused by nucleotide structure. Extending in the 3' direction interferes with the cloning sites of the M13 vector. A 3 molar excess of the above primer to target DNA gives full spontaneous priming. Greater than a 10 molar excess produces background on sequencing gels, interpreted as non-specific priming. Limiting the total amount of

-14-

primer to the specified 0.3 moles per 7 $\mu$ L reaction limits the total number of possible primed target DNA molecules. Hence, the user can use excessive amounts of DNA without depleting the nucleotide pools (up to 5 0.3 pmoles of primer at 200 number-average bases per extension is 15 pmoles of dCTP consumption for an average unknown DNA sequence, whereas a 7 $\mu$ L reaction containing 3 $\mu$ M dCTP, amounts to 21 pmoles available for consumption).

10 A preferred primer concentrate contains:

	Concentration in Final Reaction	<u>Mixture</u>
1 mM Tris/HCl (pH8).....	0.18 mM	
15 0.1 mM EDTA.....	0.018 mM	
0.24 $\mu$ M A-33 primer.....	0.3 pmoles/7 $\mu$ L	
Total Volume 5 $\mu$ L		

The third separately labeled group of 20 packages are the reaction concentrates. The reaction concentrate may be provided in the kit as four separate labeled packages, i.e. concentrate G, concentrate A, concentrate T and concentrate C. Each of the reaction concentrates will contain the same 25 chemicals at the same concentrations except the concentrations of dNTPs and ddNTPs in each will be different. The ratio of ddNTP to the dNTP have been adjusted experimentally for each concentrate to achieve balanced lane distribution of the radioactive 30 labeled dNTP. Balanced lane distribution leads to uniform appearance of photographs showing the oligonucleotide fragment distribution. This results in more precision and accuracy in reading sequence from the photographic film.

35 A very useful reaction concentrate is presented below.

-15-

		Concentration of Final Reaction	<u>Mixture</u>
	0.25 $\mu$ / $\mu$ L Tag pol.	0.5 $\mu$ /7 $\mu$ L	
5	350 $\mu$ g/ml BSA	100 $\mu$ g/ml	
	3.5 mM DTT	1 mM	
	17.5 mM Mg Acetate	5 mM	
	25 mM K Glutamate	7.1 mM	
	13.1 mM K <sub>2</sub> HPO <sub>4</sub>	3.75 mM	
10	4.4 mM KH <sub>2</sub> PO <sub>4</sub>	1.25 mM	
	0.35% Triton X-100	0.1%	
	40% Glycerol	11.4%	
	and a dNTP/ddNTP concentrate selected from Table I below.		
15.	Total Volume 2 $\mu$ L		

Table I  
dNTP Concentrations ( $\mu$ M) for Reaction Concentrates

	G	A	T	C
	daGTP	22.75	70	70
	dATP	105	17.5	105
	dTTP	52.5	52.5	52.5
	dCTP	10.5	10.5	10.5
20	ddGTP	113.75	-	-
	ddATP	-	577.5	-
	ddTTP	-	-	490
	ddCTP	-	-	231

30 The final reaction concentrations dNTP and ddNTP are shown in Table II based on the concentrations of the various chemical components in each package of the above kit. The nucleotides daGTP and dGTP are interchangeable.

-16-

Table II  
Final Reaction Concentrations in  $\mu$ M

	G	A	T	C
5	daGTP	6.5	20	20
	dATP	30	5	30
	dTTP	15	15	5
	dCTP	3	3	3
	ddGTP	32.5	-	-
10	ddATP	-	165	-
	ddTTP	-	-	140
	ddCTP	-	-	-
15	dd/d ratio	5	33	28
				22

In another embodiment of the kit the reaction concentrates G, A, T and C can also include the primer and/or the radioactive labeled dNTP.

20 In another embodiment of the kit, a single reaction concentrate may include all chemicals except the dNTPs and/or the selected ddNTP. The dNTPs and selected ddNTP could be separately packaged according to Table I. Each of these separately packaged dNTP mixtures would be added to the reaction concentrate, 25 as appropriate, at the time of use.

25 In manufacturing the reaction concentrate, the order of addition is relatively unimportant, except the enzyme should be added last. Generally the order of addition is: salts and buffers, 30 nucleotides, other reagents, glycerol and enzyme with thorough mixing upon each addition, carried out at ice temperature.

Using the kit described above, in any of the described formats, the method of sequencing would 35 generally comprise the steps of:

-17-

- a) providing a sequencing vector comprising an unknown nucleotide sequence adjacent to a 3' end of a known nucleotide chain in the vector;
- b) mixing the vector with a primer and a radioactive labeled dNTP; wherein the primer has a nucleotide sequence which is complementary to a portion of the known vector DNA sequence that is adjacent to the unknown DNA sequence;
- c) dividing the mixture b) into four equal volumes in four different vessels labeled G, A, T and C;
- d) providing a reaction mixture by:
  - i) adding to the vessel designated G a reaction concentrate according to the present invention comprising ddGTP;
  - ii) adding to the vessel designated A a reaction concentrate according to the present invention comprising ddATP;
  - iii) adding to the vessel designated T a reaction concentrate according to the present invention comprising ddTTP; and
  - iv) adding to the vessel designated C a reaction concentrate according to the present invention comprising ddCTP;
- e) incubating the four vessels G, A, T and C at a temperature of 70-74°C from 15 to 30 minutes;
- f) terminating the reactions by the addition of 3 to 5 µL of a reaction stop mixture to each of the four tubes;
- g) separating electrophoretically the oligo-nucleotides formed in e) according to length; and
- h) exposing the electrophoretic separations to a photographic film which is sensitive to the radiation generated by the radioactive dNTP to establish the nucleotide sequence of the unknown nucleotide chain.

-18-

In a typical reaction, a user provides a quality single-stranded DNA preparation, typically 1 $\mu$ g (0.4 pmoles) of M13mp18 DNA, at a concentration of >100 $\mu$ g/ml. The user combines the DNA, 5 $\mu$ L  
5 of the Primer/Buffer, water and radioactive nucleotide (usually 1 $\mu$ Ci of  $^{32}$ P or 10 $\mu$ Ci of  $^{35}$ S, alpha-labeled dCTP at >600 Ci/mM), to a total volume of 21 $\mu$ L. This DNA mixture is divided equally (5 $\mu$ L) among four tubes (preferably 0.5ml  
10 microfuge tubes), designated for the essential G-, A-, T- and C-specific reactions. These tubes are placed at 70-74°C. Then 2 $\mu$ L of the appropriate specific Reaction Concentrate is mixed into the designated tubes. The tubes are capped and the  
15 incubation continued at the 70-74°C for 30 minutes.

The reaction can be terminated with the addition of 3 $\mu$ L of reaction stop mixture. Reaction stop mixtures are well known. EPO 0 258 017 describe such mixtures as including EDTA to chelate magnesium  
20 ion, phenol, SDS or CHCl<sub>3</sub>. A useful mixture in formamide solvent is

0.02 M EDTA;  
0.07 M Tris (to neutralize EDTA);  
0.05% w/v brophenol blue; and  
25 0.02% w/v xylene cyanol.

The latter two components are anionic chromophores to provide a visual assessment of the progress of electrophoresis.

While the main advantage of using Taq pol  
30 for sequencing resides is its ability to synthesize at high temperature (minimizing structural interference by the template strand), a considerable benefit realized by this invention is its adaptability to double-stranded DNA (dsDNA) sequencing including  
35 supercoil sequencing.

-19-

The principle of double-stranded or supercoiled sequencing depends upon a fraction of supercoils being irreversibly denatured to a single-stranded state, thereby being subject to 5 primed synthesis essential to sequencing reactions.

A significant fraction of irreversibly denatured supercoils is obtained by alkaline denaturation and acid neutralization of supercoiled DNA. Normal double-stranded DNA may be similarly 10 treated, resulting in higher fraction of single-stranded DNA.

In practice, alkaline denaturation and subsequent acid neutralization of supercoil or double-stranded DNA will result in an accumulated 15 salt concentration that is high and a pH that is relatively undefined because of the relatively small, inaccurate volumes necessitated by this process. It is generally necessary to repurify the treated DNA to eliminate the salt and control the pH, since all 20 other extant sequencing protocols will not tolerate elevated salts and/or a buffered pH that would overwhelm the buffer system (usually Tris) utilized in those methods and necessary to the activity of those enzymes.

25 Phosphate buffer has the highest degree of thermal stability of any commonly used buffer. The phosphate buffer was chosen for the reaction concentrate since the pH must be controlled for the benefit of long-term storage at temperatures as low 30 as -80°C, and reaction temperatures as high as 80°C. Conversely, Tris buffer has among the lowest thermal stability of any buffer, the pK of which lowers more than two full units from 0°C to 70°C. Hence, the supercoiled sequencing reaction can be carried out 35 with the method of this invention without having to repurify the denatured DNA.

-20-

This is accomplished by:

- a) providing a double-stranded sequencing vector having a lacZ nucleotide sequence into which dsDNA having an unknown nucleotide sequence has been inserted adjacent to a 3' end of a known nucleotide chain in the vector;
- b) denaturing the vector with an alkaline solution to form ssDNA;
- c) neutralizing the denatured DNA containing solution; and
- d) carrying out steps b) through h) of the nucleotide sequencing method described hereinbefore.

In a typical example of the above sequencing method, a several  $\mu$ L volume of supercoiled DNA (in a dilute buffer, e.g. 1mM Tris, 0.1mM EDTA, pH 8) is denatured with 1 $\mu$ L of 0.1M KOH at 37°C for 2 minutes. More severe treatment is usually not necessary. On ice, the primer and the radioactive dNTP are added, and the denatured DNA is neutralized with the addition of 1 $\mu$ L of 0.2M TrisCl. Water is then added to obtain a 21 to 23  $\mu$ L volume. The final salt and pH will generally be inaccurate. However, the method of this invention can be used without further adjustments. In theory, the resulting salt system consists of a 30-50mM Tris buffer at a pH of 7-9. Upon heating this solution to 70°C for enzymatic reaction, the influence of the Tris buffer becomes negligible, since the phosphate buffer contributed by the reaction concentrate stabilizes the final pH of the entire mixture. Normal single-stranded sequencing occurs unperturbed, and supercoil sequencing occurs at the expected reduced efficiency, commensurate with the fraction of irreversibly denatured supercoils. Hence, supercoil sequencing is accomplished without repurification of the denatured supercoiled DNA, using essentially the same methodology as single strand sequencing.

-21-

The invention has been described in detail with particular reference to preferred embodiments thereof, but it will be understood that variations and modifications can be effected within the spirit 5 and scope of the invention.

10

15

20

25

30

35

-22-

We claim:

1. A nucleotide sequencing reaction concentrate comprising:

- a) sufficient thermostable polymerase to provide from 100 to 500 I.U./mL;
- 5 b) from 0.3 to 30 mM of a reducing agent selected from dithiothreitol and  $\beta$ -mercaptoethanol;
- c) sufficient phosphate buffer to maintain a pH of about 7.5;

10 d) at least 40% glycerol;

e) from 15 to 70  $\mu$ M dGTP or daGTP;

f) from 15 to 150  $\mu$ M dATP;

g) from 15 to 150  $\mu$ M dTTP;

h) from 10 to 18  $\mu$ M dCTP; and

15 i) one member selected from the group consisting of:

i) sufficient ddGTP to provide a ddGTP to dGTP or daGTP ratio of about 5;

ii) sufficient ddATP to form a ddATP to

20 dATP ratio of about 33;

iii) ddTTP in a sufficient amount to form a ddTTP to dTTP ratio of about 28; and

iv) ddCTP in a sufficient amount to form a ddCTP to dCTP ratio of about 22.

25 2. The concentrate of claim 1 wherein the thermostable polymerase is Thermus aquaticus polymerase.

3. The reaction concentrate of claim 2 also comprising:

30 a) from 175 to 700  $\mu$ g/mL of concentrate of a protein selected from the group consisting of gelatin and bovine serum albumin;

b) at least 25 mM of potassium glutamate; and

c) from 9 to 35 mM of a magnesium salt.

35 4. The reaction concentrate of claim 2 also comprising:

-23-

a) from 0.1 to 0.4% amount of a non-ionic surfactant;

b) from 0.1 to 10  $\mu\text{Ci}/\mu\text{L}$  of an alpha-labeled radioactive dNTP;

5 c) from 9 to 35 mM of a magnesium salt; and

d) from 0.1 to 1.0  $\mu\text{M}$  of an oligonucleotide primer having the nucleotide sequence:  
5' CCCAGTCACGACGTTGTAACGACGGCCAGTG 3'

5. A nucleotide sequencing reaction

10 concentrate comprising about:

a) 0.25 I.U./ $\mu\text{L}$  of Thermus aquaticus polymerase;

b) 350 micro g/mL of bovine serum albumin;

c) 3.5 mM dithiothreitol;

15 d) 17.5 mM magnesium acetate;

e) 25 mM of potassium glutamate;

f) 13.1 mM of  $\text{K}_2\text{HPO}_4$ ;

g) 4.4 mM of  $\text{KH}_2\text{PO}_4$ ;

h) 0.35% of Triton X 100 surfactant;

20 i) 40% glycerol; and one mixture selected from the group consisting of:

A) about 22.7  $\mu\text{M}$  dGTP or daGTP;  
105  $\mu\text{M}$  dATP;  
52.5  $\mu\text{M}$  of dTTP;

25 10.5  $\mu\text{M}$  of dCTP; and  
113.7  $\mu\text{M}$  of ddGTP;

B) about 70  $\mu\text{M}$  of dGTP or daGTP;  
17.5  $\mu\text{M}$  dATP;  
52.5  $\mu\text{M}$  dTTP;

30 10.5  $\mu\text{M}$  of dCTP; and  
577  $\mu\text{M}$  of ddATP;

C) about 70  $\mu\text{M}$  of dGTP or daGTP;  
105  $\mu\text{M}$  of dATP;  
17.5  $\mu\text{M}$  of dTTP;

35 10.5  $\mu\text{M}$  of dCTP; and  
490  $\mu\text{M}$  of ddTTP;

-24-

D) about 70  $\mu\text{M}$  of dGTP or daGTP;  
105  $\mu\text{M}$  of dATP;  
52.5  $\mu\text{M}$  of dTTP;  
10.5  $\mu\text{M}$  of dCTP; and  
231  $\mu\text{M}$  of ddCTP.

5

6. The reaction concentrate of claim 5  
also comprising 0.15  $\mu\text{M}$  of an oligonucleotide  
primer having the sequence:

5' CCCAGTCACGACGGTTGTAAACGACGGCCAGTG 3'

10 7. The reaction concentrate of claim 6  
also comprising 1 to 10  $\mu\text{Ci}$  of alpha-labeled  
radioactive dCTP per 2  $\mu\text{L}$  of reaction concentrate.

15 8. The reaction concentrate of claim 7  
wherein the radioactive labeled nucleotide is labeled  
with the element  $^{32}\text{P}$  or  $^{35}\text{S}$ .

9. A method of oligonucleotide sequencing  
comprising the steps of :

a) providing a sequencing vector comprising an  
unknown nucleotide sequence adjacent to a 3' end of a  
20 known nucleotide chain in the vector;

b) mixing the vector with a primer and a  
radioactive labeled dNTP; wherein the primer has a  
nucleotide sequence which is complementary to a  
portion of the known vector DNA sequence that is  
25 adjacent to the unknown DNA sequence;

c) dividing the mixture b) into four equal  
volumes in four different vessels labeled G, A, T and  
C;

d) providing a series of reaction mixtures by;  
30 i) adding to the vessel designated G a  
reaction concentrate according to claim 1 or 2  
comprising ddGTP;

ii) adding to the vessel designated A a  
reaction concentrate according to claim 1 or 2  
35 comprising ddATP;

-25-

iii) adding to the vessel designated T a reaction concentrate according to claim 1 or 2 comprising ddTTP; and

iv) adding to the vessel designated C a reaction concentrate according to claim 1 or 2 comprising ddCTP;

5 e) incubating the four vessels G, A, T and C at a temperature of 70-74°C from 15 to 30 minutes;

f) terminating the reactions by the addition of 10 3 to 5 µL of a reaction stop mixture to each of the four tubes;

g) separating electrophoretically the oligo-nucleotides formed in e) according to length; and

15 h) exposing each gel to a photographic film that is sensitive to the radiation generated by the radioactive dNTP to establish the nucleotide sequence of the unknown nucleotide chain.

10. A method according to claim 9 wherein each reaction concentrate G, T, A and C already 20 includes a radioactive dNTP thereby obviating the need to include said radioactive dNTP in step b) of claim 9.

11. The method of claim 9 wherein each reaction concentrate, G, T, A and C, also includes 25 the primer thereby obviating the need to include said primer in step b) of claim 9.

12. The method of claim 9 wherein the sequencing vector has a lacZ nucleotide sequence.

13. The method of claim 9 wherein the 30 sequencing vector is selected from the group consisting of M13mpl8 DNA, pUC, pIBI and Bluescript.

14. The method of claim 9 wherein the primer has the nucleotide sequence:

5' CCCAGTCACGACGTTGTAAACGACGGCCAGTG 3'

35 15. The method of claim 9 or 11 in which the sequencing vector is M13mpl8 and the reaction

-26-

concentrates have compositions according to claim 2,  
3, 4 or 5.

16. A nucleotide sequencing kit comprising:

- a) a control sequencing vector having a lacZ  
5 nucleotide sequence and 1 $\mu$ L of alpha-labeled  
radioactive dCTP in a first container;
- b) an oligonucleotide primer having the  
sequence:

5' CCCAGTCACGACGTTGTAAAACGACGGCCAGTG 3'

10 in a second container; and

- c) a reaction concentrate according to claim 2,  
3 or 5 in a second set of containers.

17. A nucleotide sequencing kit comprising  
an oligonucleotide primer having the sequence:

15 5' CCCAGTCACGACGTTGTAAAACGACGGCCAGTG 3'  
in a first container and a reaction concentrate  
according to claim 2, 3 or 5 in a second set of  
containers.

18. A nucleotide sequencing kit comprising  
20 a reaction concentrate according to claim 6, 7 or 8.

19. An oligonucleotide having the  
nucleotide sequence:

5' CCCAGTCCACGACGTTGTAAAACGACGGCCAGTG 3'

20. A method of sequencing double-stranded  
25 (dsDNA), comprising the steps of:

- a) providing a double-stranded sequencing  
vector having a lacZ nucleotide sequence into which  
dsDNA having an unknown nucleotide sequence has been  
inserted adjacent to a 3' end of a known nucleotide  
30 chain in the vector;
- b) denaturing the vector with an alkaline  
solution to form ssDNA;
- c) neutralizing the denatured DNA containing  
solution;
- 35 d) mixing the vector with a primer, a magnesium  
salt and a radioactive labeled dNTP; wherein the

-27-

primer has a nucleotide sequence which is complementary to at least a portion of known vector DNA sequence that is adjacent to the unknown DNA sequence;

5 e) dividing the mixture d) into four equal volumes in four different vessels labeled G, A, T and C; and

f) providing a series of reaction mixtures by:

i) adding to the vessel designated G a reaction concentrate according to claim 1 or 2

10 comprising ddGTP;

ii) adding to the vessel designated A a reaction concentrate according to claim 1 or 2 comprising ddATP;

15 iii) adding to the vessel designated T a reaction concentrate according to claim 1 or 2 comprising ddTTP;

iv) adding to the vessel designated C a reaction concentrate according to claim 1 or 2 comprising ddCTP;

20 g) incubating the four vessels G, A, T and C at a temperature of 70-74°C from 15 to 30 minutes;

h) terminating the reactions by the addition of 3 to 5 µL of a reaction stop mixture to each of the four tubes;

25 i) separating electrophoretically the formed oligonucleotides according to length; and

j) exposing each gel to a photographic film which is sensitive to the radiation generated by the radioactive dNTP to establish the sequence of the

30 supercoiled DNA.

**INTERNATIONAL SEARCH REPORT**

International Application No. PCT/US 90/00440

**I. CLASSIFICATION OF SUBJECT MATTER** (if several classification symbols apply, indicate all)<sup>6</sup>

According to International Patent Classification (IPC) or to both National Classification and IPC  
IPC5: C 12 Q 1/68, C 12 N 9/12

**II. FIELDS SEARCHED**

Classification System	Minimum Documentation Searched <sup>7</sup>	
	Classification Symbols	
IPC5	C 12 Q; C 12 N	
Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in Fields Searched <sup>8</sup>		

**III. DOCUMENTS CONSIDERED TO BE RELEVANT<sup>9</sup>**

Category <sup>10</sup>	Citation of Document <sup>11</sup> with indication, where appropriate, of the relevant passages <sup>12</sup>	Relevant to Claim No. <sup>13</sup>
X	EP, A2, 0258017 (CETUS CORPORATION) 2 March 1988, see page 30, line 50 - page 31, line 10  --	1-5, 9- 13, 20
A	WO, A1, 8805470 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 28 July 1988, see page 44, line 22 - line 34  --	1-20
A	Proc. Natl. Acad. Sci., Vol. 85, 1988, (USA) Michael A Innis et al: "DNA sequencing with Thermus aquaticus DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA ", see page 9436 - page 9440  -- -----	1-20

\* Special categories of cited documents:<sup>10</sup>

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

**V. CERTIFICATION**

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
---	---

7th May 1990

- 5. 06. 90

International Searching Authority

Signature of Authorized Officer

EUROPEAN PATENT OFFICE

F.W. HECK

ANNEX TO THE INTERNATIONAL SEARCH REPORT  
ON INTERNATIONAL PATENT APPLICATION NO.PCT/US 90/00440

SA 34668

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.  
The members are as contained in the European Patent Office EDP file on 07/05/90  
The European Patent office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date		Patent family member(s)	Publication date
EP-A2- 0258017	02/03/88	AU-D- JP-A-	7729687 63102677	19/05/88 07/05/88
WO-A1- 8805470	28/07/88	AU-D- EP-A- JP-A- US-A-	1022488 0265293 63237798 4795699	21/07/88 27/04/88 04/10/88 03/01/89

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

EPO FORM P0479